Increasing hazard rate of mixtures for natural exponential families

Shaul K. Bar-Lev* and Gérard Letac†

April 27, 2010

Abstract

Hazard rates play an important role in various areas, e.g., reliability theory, survival analysis, biostatistics, queueing theory and actuarial studies. Mixtures of distributions are also of a great preeminence in such areas as most populations of components are indeed heterogeneous. In this study we present a sufficient condition for mixtures of two elements of the same natural exponential family (NEF) to have an increasing hazard rate. We then apply this condition to some classical NEF's having either quadratic, or cubic variance functions (VF) and others as well. A particular attention is devoted to the hyperbolic cosine NEF having a quadratic VF, the Ressel NEF having a cubic VF and to the Kummer distributions of type 2 NEF. The application of such a sufficient condition is quite intricate and cumbersome, in particular when applied to the latter three NEF's. Various lemmas and propositions are needed then to verify this condition for these NEF's.

Key words: Natural exponential families; mixtures; variance functions; quadratic variance functions; cubic variance functions; hyperbolic cosine NEF; Ressel NEF; Kummer type 2 NEF.

1 Introduction

Hazard rates (also called failure rates) play an important role in various areas, e.g., reliability theory, queuing models, survival analysis and actuarial studies. Mixtures of distributions are also of a great preeminence in such areas as most populations of components are indeed heterogeneous. A comprehensive list of references on the behavior of hazard rates for mixtures of distributions can be found in the monograph by Shaked and Shanthikumar (2007) and the references cited therein and also in Block et al. (2003).

*Department of Statistics, University of Haifa, Haifa 31905, Israel (barlev@stat.haifa.ac.il)

†Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062 Toulouse, France (gerard.letac@alsatis.net)

In this study we present (Proposition 2) a sufficient condition for mixtures of two elements of the same natural exponential family (NEF) to have an increasing hazard rate. We then apply this condition to some classical absolutely continuous NEF's having either quadratic or cubic variance functions (VF's) (c.f., Morris, 1982, Letac and Mora, 1990) and others NEF's as well. A particular attention is devoted to the hyperbolic cosine NEF having a quadratic VF, the Ressel NEF having a cubic VF and the Kummer distributions of type 2 NEF. The application of such a sufficient condition can be intricate, in particular when applied to the latter three NEF's. Various lemmas and propositions are then needed to verify this condition for such NEF's. Accordingly, we dedicate Sections 4, 5 and 6, respectively, for these three NEF's. In Section 3 we consider the rather easy application of the sufficient condition to three NEF's having either quadratic or cubic VF's, namely, the normal, gamma and the inverse Gaussian NEF's. Our sufficient condition stems from the following seminal result by Glaser (1980):

Proposition 1 *Suppose that the probability density $s(x)$, concentrated on the interval (a,∞) (with $-\infty \leq a < \infty$), is positive such that $-b(x) = \log s(x)$ is concave. Then the mapping $x \mapsto \log \int_x^\infty s(t)dt$ is concave on (a,∞) and the hazard function $h(x) = s(x)/\int_x^\infty s(t)dt$ is increasing.*

(Ron Glaser observes for the one line proof that since $b'(x)$ is nondecreasing, one has $(1/h)'(x) = \int_x^\infty e^{b(x)-b(t)}(b'(t)-b'(x))dt \leq 0$). If the probability density is $s = e^{-b}$ the fact that b is convex is by Proposition 1 a sufficient condition for having h increasing but *not* a necessary one: see the remark in Section 2 introducing the Glaser set as well the Jorgensen and the Karlin sets, or consider the density $s_2 = e^{-b_2}$ in Section 5 for which h in increasing and b_2 is not convex.

Our sufficient condition for a mixture of two members in the same NEF to have an increasing hazard rate is as follows: Suppose that the NEF is written as

$$\{e^{-\lambda x - k(\lambda) - b(x)}\mathbf{1}_{(a,\infty)}(x)dx,\ \lambda \in \Lambda\},$$

where Λ is an interval and $-\infty \leq a$. Suppose also that $b''(x) \geq 0$ for all $x > a$ and denote $T(x) = 1/\sqrt{b''(x)}$. We show in Proposition 2 that if there exists $c > 0$ and $d \in \mathbb{R}$ such that the inequality $cT(x) \leq \cosh(cx + d)$ holds for all $x > a$, we then can find pairs λ_1 and $\lambda_2 = \lambda_1 + 2c$ in Λ and a mixing coefficient $p \in (0,1)$ such that the mixture density

$$\left(pe^{-\lambda_1 x - k(\lambda_1)} + (1-p)e^{-\lambda_2 x - k(\lambda_2)}\right)e^{-b(x)}$$

has an increasing hazard rate. This simple condition relies on the fact that the mixture is employed with two elements of the same NEF. However, this two element mixture result is apparently not extendable to a more multi-element mixture situation.

2 A sufficient condition for mixtures of members of the same NEF to have an increasing hazard rate

Consider an absolutely continuous NEF concentrated on (a, ∞) with $-\infty \leq a < \infty$, and generated by a locally integrable function s on (a, ∞) such that

$$L(\lambda) = e^{k(\lambda)} = \int_a^\infty e^{-\lambda x} s(x) dx, \tag{1}$$

the Laplace transform (LT) of $s(x)$, exists on a nonempty open interval Λ. The corresponding NEF is then given by the set of probability densities on (a, ∞) of the form

$$\{\exp\{-\lambda x - k(\lambda)\} s(x) dx,\ \lambda \in \Lambda\}. \tag{2}$$

Let $\nu(d\lambda)$ be a probability on Λ and suppose that the function on (a, ∞) defined by

$$R(x) = \int_\Lambda e^{-\lambda x} \frac{\nu(d\lambda)}{L(\lambda)} \tag{3}$$

exists. Thus, $s(x)R(x)dx$ is a probability density on (a, ∞) and it is a mixture of the elements of the NEF. This probability density has the hazard rate

$$h_\nu(x) = \frac{s(x)R(x)}{\int_x^\infty s(t)R(t)dt}. \tag{4}$$

Proposition 1 shows that $h_\nu(x)$ is increasing if $x \mapsto \log(s(x)R(x))$ is concave, or, equivalently, if $s(x) > 0$ for all $x > a$, if $s''(x)$ exists and if on (a, ∞) one has

$$\frac{s''(x)s(x) - (s'(x))^2}{s^2(x)} + \frac{R''(x)R(x) - (R'(x))^2}{R^2(x)} \leq 0. \tag{5}$$

We now have the following proposition when ν is a mixture of two Dirac measures.

Proposition 2 *Consider the special case of the hazard rate h_ν in (4) with*

$$\nu = p\delta_{\lambda_1} + (1-p)\delta_{\lambda_2}, \tag{6}$$

where $p \in (0,1)$ and $\lambda_1 < \lambda_2$ with λ_1 and λ_2 in Λ. Assume that on (a, ∞), $s(x) > 0$, $-b(x) = \log s(x)$ is concave and that $s''(x)$ exists and define

$$T(x) = 1/\sqrt{b''(x)} \tag{7}$$

$$p_1 = pe^{-k(\lambda_1)},\ p_2 = (1-p)e^{-k(\lambda_2)},\ c = \frac{\lambda_2 - \lambda_1}{2} \text{ and } d = \log\sqrt{p_1/p_2}. \tag{8}$$

Then the hazard rate (4) with ν as in (6) is increasing if for all $x > a$

$$cT(x) \leq \cosh(cx + d). \tag{9}$$

Proof. The proof is a straightforward application of (5). Indeed, for ν in (6) and R defined by (3),

$$R(x) = pe^{-k(\lambda_1)}e^{-\lambda_1 x} + (1-p)e^{-k(\lambda_2)}e^{-\lambda_2 x} = p_1 e^{-\lambda_1 x} + p_2 e^{-\lambda_2 x},$$

implying that $R''(x)R(x) - (R'(x))^2 = p_1 p_2 (\lambda_2 - \lambda_1)^2 e^{-(\lambda_1+\lambda_2)x}$. Accordingly, the inequality (5) becomes for this particular case

$$p_1 p_2 (\lambda_2 - \lambda_1)^2 e^{-(\lambda_1+\lambda_2)x} \leq (p_1 e^{-\lambda_1 x} + p_2 e^{-\lambda_2 x})^2 \left(-\frac{s'(x)}{s(x)}\right)', \; x > a. \quad (10)$$

Since $\log s(x)$ is concave, $(-s'(x)/s(x))' \geq 0$ on (a, ∞), so that $T(x)$ in (7) is well defined. Thus, with the notations (8), the inequality (10) is equivalent to

$$2cT(x) = T(x)(\lambda_2 - \lambda_1) \leq \sqrt{\frac{p_1}{p_2}} e^{-\frac{\lambda_1 - \lambda_2}{2}x} + \sqrt{\frac{p_2}{p_1}} e^{-\frac{\lambda_2 - \lambda_1}{2}x} = 2\cosh(cx + d).$$

which is (9). ■

In the next sections, we are going to consider a number of absolutely continuous NEF's on the real line, generated by a density s, and check for each of them whether T exists or not, that is whether s is log concave or not. When T exists we will have to discover which (c, d) with $c > 0$ are such that (9) holds for all x. As we shall see, for some NEF's such that s is log concave it may occur that (9) does not hold for any (c, d).

The system (8) of equalities links the three parameters $(\lambda_1, \lambda_2, p)$ with the two parameters (c, d). Suppose that we are given a pair (c, d) satisfying (9), we therefore may choose arbitrarily the mean $\lambda = \frac{\lambda_1+\lambda_2}{2}$ in Λ such that $\lambda_1 = \lambda - c$ and $\lambda_2 = \lambda + c$ are in Λ. Having done this choice of λ, the value of of the mixing coefficient p in (8) can be determined exactly as

$$p = \frac{e^d L(\lambda - c)}{e^d L(\lambda - c) + e^{-d} L(\lambda + c)}, \quad (11)$$

where L is the Laplace transform (LT) of the generating density s (recall that s is not necessarily a probability). We note, however, that the LT is not always expressible in terms of simple functions but rather in terms of transcendental or implicit functions, in which case a numerical search is then needed to find the (c, d) interval on which the appropriate mixture density possesses an increasing hazard rate. As this paper is rather theoretical, we do not intend to pursue such a numerical search.

Remarks on the Jorgensen, Karlin and Glaser sets. Given a density s on (a, ∞) with LT (1) such that Λ is not empty, the set $J(s)$ of $\alpha \geq 0$ such that L^α is still a LT of a positive measure μ_α is called the Jorgensen set of s (see for instance Letac and Mora (1990) and the references cited therein). By definition, $J(s)$ is a closed additive semigroup. Note that s generates an NEF of infinitely divisible distributions if and only if $J(s) = [0, \infty)$. If not, $J(s)$ can

be complicated. For instance, a consequence of the short and elegant paper by Ben Salah and Masmoudi (2010) is that $J(s) = [1, \infty)$ if

$$s(x) = \frac{1}{4}e^{-x}\mathbf{1}_{(0,\infty)}(x) + \frac{3}{4}e^{-x+1}\mathbf{1}_{(1,\infty)}(x).$$

Note also that μ_α could have a continuous singular part for some small $\alpha \in J(s)$ although appropriate examples are rather complicated. If $J^*(s) \subset J(s)$ is the set of α such that $\mu_\alpha(dx) = s_\alpha(x)dx$ has density we trivially have that $J^*(s) + J(s) \subset J^*(s)$, since the convolution of a measure with a density with any measure has a density.

Consider now a non trivial result due to Karlin and Proschan (1960) (see also Karlin, 1968, p. 152 and Barlow and Proshan, 1965, p. 100) which says that if s and ℓ are probability densities with increasing hazard rate, then the convolution $s * \ell$ has the same property. Therefore let us introduce the Karlin set $K(s)$ of $\alpha \in J^*(s)$ such that s_α has an increasing hazard rate. The above property shows that $K(s)$ is a closed additive subsemigroup of $J^*(s)$. For instance if $s(x) = e^{-x}\mathbf{1}_{(0,\infty)}(x)$, it is a simple exercise to see that $K(s) = [1, \infty)$.

Finally, consider the Glaser set $G(s)$ of s which is the set of α in the Karlin set $K(s)$ for which the conditions of Proposition 1 are met, that is such that if $s_\alpha = e^{-b_\alpha}$ then b_α is convex. Although in many cases $G(s)$ coincides with $K(s)$ the Glaser set $G(s)$ is not a semigroup. Indeed, we shall face with an example in Section 4 relating to the Ressel distribution s_1 where $G(s)$ is a bounded interval; thus distinct from the semi group $K(s)$. Usually the Karlin set is more difficult to find than the Glaser set.

3 Applications related to NEF's with quadratic or cubic VF's (normal, gamma and inverse Gaussian NEF's)

As already noted in the introduction, quadratic VF's include six NEF's of which only three have densities: Normal, gamma and hyperbolic cosine (c.f. Morris, 1982). Cubic VF's include also six NEF's of which only two have densities: Inverse Gaussian and Ressel (c.f. Letac and Mora, 1990). The set of our examples will include all of the five absolutely continuous NEF's having either quadratic or cubic VF's and also another one, the Kummer distributions of type 2 NEF. The present section deals with the normal, gamma and inverse Gaussian NEF's and Sections 4, 5 and 6 consider the three other ones. In what follows and whenever feasible, we provide, for each of the examples, with their respective VF (V, Ω), where V is the VF corresponding to (2) and Ω is the domain of means.

Example 1: *The normal NEF*
The normal NEF has a constant VF, i.e. $(V, \Omega) = (\sigma^2, \mathbb{R})$. For a fixed

standard deviation σ, the generating density is

$$s(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \; a = -\infty.$$

Trivially here the Glaser set $G(s)$ is $(0, \infty)$. This leads to $T(x) = \sigma$ and $k(\lambda) = \frac{\sigma^2 \lambda^2}{2}$. The inequality (9) is fulfilled for any $x \in \mathbb{R}$ if and only if $c\sigma \leq 1$, or equivalently, if $|\lambda_1 - \lambda_2| \leq 2/\sigma$, a result that was already obtained by Block *et al.* (2005).

Example 2: *The gamma NEF*

The gamma NEF, concentrated on $(0, \infty)$, has a VF $V(\mu) = \alpha^{-1}\mu^2$ and $\Omega = \mathbb{R}^+$, where α and μ are, respectively, the shape and mean parameters. For a fixed shape parameter $\alpha > 0$, the generating measure is

$$s_\alpha(x)dx = \frac{x^{\alpha-1}}{\Gamma(\alpha)} \mathbf{1}_{(0,\infty)}(x)dx.$$

We now consider three exhausted cases relating to the values of the parameter α: $\alpha = 1$ (the exponential case), $\alpha < 1$ and $\alpha > 1$. These observations imply that the Glaser set $G(s_1)$ is $[1, \infty)$.

1. $\alpha = 1$. Here, $s(x) \equiv 1$ so that for any ν the function $\log R$ is convex and the inequality (9) (as well as (5)) cannot be fulfilled unless ν is concentrated on one point.
2. $\alpha < 1$. Since for this case both $\log R$ and $\log s$ are convex, the inequalities (5) or (9) cannot be fulfilled.
3. $\alpha > 1$. Here, $T(x) = x/\sqrt{\alpha - 1}$ and we have the following proposition.

Proposition 3 *For $\alpha > 1$, the probability density*

$$f(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1}(p\lambda_1^\alpha e^{-\lambda_1 x} + (1-p)\lambda_2^\alpha e^{-\lambda_2 x}), x > 0,$$

where $\lambda_1 < \lambda_2$, has an increasing hazard rate $h(x) = f(x)/\int_x^\infty f(t)dt$ if

$$\frac{\lambda_2}{\lambda_1} \leq \left(\frac{p}{1-p}\right)^{1/\alpha} e^{-d_0/2\alpha},$$

where

$$d_0 = \log \frac{1 + \sqrt{\alpha}}{\sqrt{\alpha - 1}} - \frac{2\alpha}{(1 + \sqrt{\alpha})}.$$

Example. For $\alpha = 2$ the result specializes to the following: since $d_0 = \log 2 - 2$

and since $e^{d_0/2} = e/\sqrt{2} = 1.92...$ we can claim that the following mixing of two gamma densities

$$f(x) = p\lambda_1^2 x e^{-\lambda_1 x} + (1-p)\lambda_2^2 x e^{-\lambda_2 x}$$

where $\lambda_1 < \lambda_2$ has an increasing hazard rate if $\left(\frac{\lambda_2}{\lambda_1}\right)^2 \leq \frac{p}{1-p}(1.92..)$. For instance choosing $\lambda_2 = 2\lambda_1$ imposes a heavy weight p on λ_1, namely $.675 < p < 1$.

Proof. Since $c > 0$ for studying the inequality (9) we write $t = cx + d$. Thus

(9) becomes: for all $t > -d$,

$$\frac{t-d}{\sqrt{\alpha-1}} \leq \cosh t. \tag{12}$$

For a fixed $\alpha > 1$, we determine the set of d values such that (12) holds. Since $t \mapsto \cosh t$ is a convex function, we look for the point $(t_0, \cosh t_0)$ such that the tangent to the curve cosh has slope $1/\sqrt{\alpha-1}$. Thus,

$$\sinh t_0 = \frac{1}{\sqrt{\alpha-1}},\ t_0 = \log\frac{1+\sqrt{\alpha}}{\sqrt{\alpha-1}} \text{ and } \cosh t_0 = \frac{2\alpha}{(1+\sqrt{\alpha})\sqrt{\alpha-1}}.$$

The equation of this tangent is $y = (t - d_0)/\sqrt{\alpha-1}$, where d_0 is such that this line goes through the point $(t_0, \cosh t_0)$. This implies that

$$d_0 = t_0 - (\cosh t_0)\sqrt{\alpha-1} = \log\frac{1+\sqrt{\alpha}}{\sqrt{\alpha-1}} - \frac{2\alpha}{(1+\sqrt{\alpha})}. \tag{13}$$

Such results show that $t/\sqrt{\alpha-1} \leq \cosh(t+d)$ for all $t > 0$ if and only if $d \geq d_0$. The application of this fact is that (10) holds for all $x > 0$ if and only if $\sqrt{p_1/p_2} \geq e^{d_0}$, or, equivalently, if

$$\frac{\lambda_1}{\lambda_2} \geq \left(\frac{1-p}{p}\right)^{1/\alpha} e^{d_0/2\alpha}.$$

■

Example 3: *The inverse Gaussian NEF* The inverse Gaussian NEF has a VF $V(\mu) = \alpha^{-2}\mu^3$ with $\Omega = \mathbb{R}^+$, where $\alpha > 0$ and μ is the mean parameter. Here $a = 0$ and for a fixed $\alpha > 0$, the corresponding NEF is generated by

$$s(x) = \frac{\alpha}{\sqrt{2\pi}} x^{-3/2} e^{-\frac{\alpha^2}{2x}}.$$

This implies that $(-s'(x)/s(x))' = (3x - 2\alpha^2)/2x^3$, which is not a positive function. Thus Proposition 2 is not applicable.

4 The hyperbolic cosine NEF

The hyperbolic cosine NEF $\mathcal{F}_\alpha$ has a VF $V(\mu) = \frac{\mu^2}{\alpha} + \alpha$ with $\Omega = \mathbb{R}$, where $\alpha > 0$. The convex support of $\mathcal{F}_\alpha$ is $\mathbb{R}$ (i.e., $a = -\infty$). The generating measure of $\mathcal{F}_\alpha$ is

$$\mu_\alpha(dx) = \frac{2^{\alpha-2}}{\pi}\left|\Gamma(\frac{\alpha+ix}{2})\right|^2\frac{dx}{\Gamma(\alpha)}.$$

(See Morris 1982, for details). Moreover, its LT is defined on $\Lambda = (-\frac{\pi}{2}, \frac{\pi}{2})$ by $L_\alpha(\lambda) = (\cos\lambda)^{-\alpha}$ (since it is 1 for $\lambda = 0$, this shows the non obvious fact that μ_α is a probability). The hyperbolic cosine μ_1 distribution is more known as the hyperbolic secant (hereafter, HS distribution or HS NEF). Various probabilistic properties of the HS distribution have been derived, though it is rarely used in applied statistics, probably due to its intricate structure. Although this distribution is not used much in applications, it does, however, have two curious features: Like the normal distribution, the density of μ_1 is proportional to its characteristic function; the sample mean and median are, asymptotically, equally efficient. A probabilistic interpretation of μ_1 is available: consider a standard complex Brownian motion $Z = X + iY$ with $Z(0) = 0$ and the hitting time T of the set $\{x + iy\ ; |y| \geq \pi/2\}$. Then $X(T) \sim \mu_1$: to see this, consider the process $M(t) = \exp sZ(t)$. Since $z \mapsto e^{sz}$ is analytic, it is harmonic, M is a martingale and $\mathbb{E}(M(T)) = 1$ gives the desired result. A newsworthy statistical analysis and data fitness can be found in Smyth (1994) and recently in Sibuya (2006) (a complete English version of the latter paper is available by corresponding the author).

Denote by $s_\alpha = e^{-b_\alpha}$ the density of μ_α. The fact that the function b_α is convex if and only if $\alpha \geq 1$ has been proved by Shanbhag (1979). We give a different proof in the following proposition:

Proposition 4 *The function b_α is convex if and only if $\alpha \geq 1$ (in other terms the Glaser set $G(s_1)$ is $[1, \infty)$). More specifically for $\alpha > 1$ we have*

$$\frac{1}{s_\alpha(x)} = (\alpha - 1)\int_{-\pi/2}^{\pi/2} e^{xu}(\cos u)^{\alpha-2}\,du. \tag{14}$$

and for $\alpha < 1$ the function b''_α is negative in the interval

$$\left(\alpha\sqrt{\frac{2+\alpha}{2-\alpha}}\ ,\ \frac{2+\alpha}{\sqrt{3}}\right).$$

Proof. Formula (14) is the particular case $\nu = \alpha - 1 > 0$ and $a = ix$ of the classical formula

$$\int_0^{\pi/2}(\cos u)^{\nu-1}\cos au\,du = \frac{\pi}{2^\nu\nu B(\frac{\nu+1+a}{2}, \frac{\nu+1-a}{2})} \tag{15}$$

which can be found in Gradshteyn and Ryzhik (1980), page 372, 3.631 formula 9. Now (14) shows that $1/s_\alpha$ is the Laplace transform of the positive measure

$$(\alpha - 1)(\cos u)^{\alpha-2}\mathbf{1}_{(-\pi/2,\pi/2)}(u)\,du \tag{16}$$

which implies that the function $b_\alpha = -\log s_\alpha$ is strictly convex for $\alpha > 1$. For $\alpha = 1$ we can see that $b_1 = -\log s_1$ is convex by the same trick since $1/s_1 = 2\cosh\frac{\pi x}{2}$ is the Laplace transform of the positive measure $\delta_{-\pi/2} + \delta_{\pi/2}$. Note that is the weak limit of (16) when $\alpha \to 1$.

Suppose now that $0 < \alpha < 1$. We use the digamma function $\psi = \Gamma'/\Gamma$ and its derivative. If z is a complex number with positive real part :

$$\psi'(z) = \sum_{n=0}^{\infty} \frac{1}{(n+z)^2} \tag{17}$$

A easy calculation leads to

$$b''_\alpha(x) = \frac{1}{4}\psi'(\frac{\alpha+ix}{2}) + \frac{1}{4}\psi'(\frac{\alpha-ix}{2}) = \frac{1}{4}\varphi_{\alpha/2}(\frac{x^2}{4})$$

where for $t > 0$ we define

$$\varphi_c(t) = \frac{c^2-t}{(c^2+t)^2} + \sum_{n=1}^{\infty} \frac{(n+c)^2-t}{((n+c)^2+t)^2}.$$

For showing that for $c < 1/2$ the function $t \mapsto \varphi_c(t)$ is negative on some interval observe that for fixed $t > 0$ the function on $(0,\infty)$ defined by $u \mapsto \frac{u-t}{(u+t)^2}$ is decreasing when $u > 3t$. As a consequence if $(1+c)^2 > 3t$ we can write the majorization of the sum of a series by an integral

$$\varphi_c(t) < \frac{c^2-t}{(c^2+t)^2} + \int_0^\infty \frac{(v+c)^2-t}{((v+c)^2+t)^2} dv = \frac{c^2-t}{(c^2+t)^2} + \frac{c}{c^2+t}$$

(here we have used $\frac{v^2-t}{(v^2+t)^2} = -\frac{d}{dv}\frac{v}{v^2+t}$). This shows that $\varphi_c(t) < 0$ when $\frac{c^2(1+c)}{1-c} \le t \le \frac{(1+c)^2}{3}$. Since $c < 1/2$ we have $\frac{(1+c)^2}{3} - \frac{c^2(1+c)}{1-c} = \frac{1+c}{3(1-c)}(1-4c^2) > 0$ and this interval is not empty. Replacing c by $\alpha/2$ and t by $x^2/4$ we get that b''_α is negative in the interval indicated in the statement of the proposition.

■

For part $0 < \alpha < 1$ our proof of Proposition 4 is elementary. For part $\alpha \ge 1$ our proof is based on the formula 15 and Laplace transforms. The compact and ingenious Shanbhag's proof rather relies on Fourier transforms through the formula $b''_\alpha(x) = \int_{-\infty}^{\infty} e^{itx/2} d_\alpha(t)dt$ where

$$d_\alpha(t) = \frac{t}{2\sinh\frac{t}{2}} e^{\frac{|t|}{2}(1-\alpha)}.$$

This formula is derived from an integral formula for b_α which can be found in Zolotarev (1967) and which is obtained from the Lévy measure of the infinitely divisible distribution of $\log X$ when X is $\gamma_{\alpha/2}$ distributed. If $\alpha \ge 1$ the function d_α is an integrable characteristic function (corresponding to a Cauchy distribution with parameter $(\alpha-1)/2$ convoluted with the density $\pi/2(\cosh \pi x)^2$)).

The Fourier inversion formula shows that $b''_\alpha(x) \geq 0$ for all x. If $0 < \alpha < 1$ we have $d_\alpha(x) > 1$ around zero and d_α cannot be a characteristic function. This prevents b''_α to be positive by a careful but standard reasoning using again the Fourier inversion and this concludes the Shanbhag's proof.

Since the functions b_α and $T = 1/\sqrt{b''_\alpha}$ are not simple when α is not an integer we therefore emphasize the analysis of the respective mixtures for the two cases: $\alpha = 1$ and $\alpha = 2$. In principle an analysis similar to the case $\alpha = 2$ below could be also performed for $\alpha = 3, 4, \ldots$ but the case $\alpha = 2$ is creative enough to let us think that higher cases are difficult.

The hyperbolic case $\alpha = 1$: The most popular member of the $\mathcal{F}_\alpha$'s is related to this case. As mentioned in the proof of Proposition 4 the corresponding density for $\alpha = 1$ is

$$s_1(x) = \frac{1}{2\cosh\frac{\pi x}{2}}.$$

Thus the above results are applicable to this $\mathcal{F}_1$. More specifically

$$b''_1(x) = \left(-\frac{s'_1(x)}{s_1(x)}\right)' = (\frac{\pi}{2})^2 \frac{1}{\cosh^2\frac{\pi x}{2}} > 0,$$

and thus $T(x) = \frac{2}{\pi}\cosh\frac{\pi x}{2}$. In order to study the inequality (9) for this particular case, we use the following lemma.

Lemma 5 *Let a and u be positive numbers and v be a real number. Then, the following inequality*

$$a\cosh x \leq \cosh(ux + v)$$

holds for all real x if and only if $a \in (0, 1]$, $u \geq 1$ and $|v| \leq v_0 = v_0(a, u)$, where

$$v_0 = u\log(\frac{A}{a} + \frac{uB}{a}) - \log(A + B)$$

with the notation $A = \sqrt{\frac{u^2-a^2}{u^2-1}}$ and $B = \sqrt{\frac{1-a^2}{u^2-1}}$.

Proof. $\Rightarrow$. Letting $x \to \infty$, we have that $a\cosh x \sim ae^x$ and $\cosh(ux+v) \sim e^{ux}$,

implying that $u \geq 1$. Letting $x = -v/u$ shows that $a\cosh(-u/v) \leq 1$ and thus $a \leq 1$. In the sequel we assume that $u > 1$ and treat the case $u = 1$ separately after. Now, we introduce the two positive numbers x_0 and v_0 such that the two curves $x \mapsto a\cosh x$ and $x \mapsto \cosh(ux - v_0)$ are tangent on a point of the abscissa v_0. Thus, they satisfy the two equations:

$$a\cosh x = \cosh(ux_0 - v_0) \text{ and } \ a\sinh x = u\sinh(ux_0 - v_0).$$

Squaring these two equations and using the fact that $\cosh^2 t - \sinh^2 t = 1$, we get a linear system in $\cosh^2 x_0$ and $\cosh^2(ux_0 - v_0)$, whose solution is

$$\cosh^2 x_0 = \frac{A^2}{a^2} \text{ and } \cosh^2(ux_0 - v_0) = A^2.$$

Since $t \geq 0$ and $y = \cosh t$, it follows that $t = \log(y + \sqrt{y^2 - 1}$. Thus, $x_0 = \log\left(\frac{A}{a} + \frac{uB}{a}\right)$ and $v_0 = u \log\left(\frac{A}{a} + \frac{uB}{a}\right) - \log(A + B)$ (note that $ux_0 - v_0 \geq 0$). To end the proof of $\Rightarrow$ we show that $a \cosh x \leq \cosh(ux + v)$ for all real x would imply that $|v| \leq v_0$. Since the function $v \mapsto \cosh(ux_0 - v) - a\cosh x_0$ is decreasing on the interval $(-\infty, ux_0)$ and is zero on v_0 (which belongs to this interval), we get that $v < v_0$ when $\cosh(ux_0 - v) - a \cosh x_0 \geq 0$. Similarly, because of the symmetry of $\cosh t$ we have $-v_0 \leq v$.

We now prove $\Leftarrow$. Assume that $a \leq 1 < u$ and $|v| \leq v_0$. Denoting $f(x) = \cosh(ux - v) - a\cosh x$, then since $f(x_0) = f'(x_0) = 0$, the Taylor formula gives

$$f(x) = \int_{x_0}^{x} (x - t) f''(t) dt. \tag{18}$$

We use the latter formula to show that $f(x) > 0$ for $x > x_0$. Note that since $f''(x) > f(x)$, then $f''(x_0) > 0$ and (18) implies that $f(x) > 0$ on some interval (x_0, x_1). Now suppose that there exists $x_2 > x_0$ such that $f(x_2) = 0$. Without loss of generality we assume that $f(x) > 0$ on (x_0, x_2). Thus $f''(x) > 0$ on (x_0, x_2). Since by (18), $f(x_2) = 0$ is impossible, we obtain that $f(x) > 0$ for all $x > x_0$. To prove that $f(x) > 0$ for all $x < x_0$ is similar.

We now consider the particular case $u = 1$. The inequality $a \cosh x \leq \cosh(x + v)$ is equivalent to $e^{2x}(e^v - a) \geq a - e^{-v}$. By letting $x \to \pm\infty$, it can be easily seen that the latter inequality holds for all x if and only if $|v| \leq -\log a$. ■

We do not apply the full strength of this lemma for our problem, but instead study when the inequality

$$\frac{2c}{\pi} \cosh \frac{\pi x}{2} \leq \cosh(cx + d),$$

holds for all x.

To fit with the notations of the latter lemma, denote $t = \frac{\pi x}{2}$ which leads to $\frac{2c}{\pi} \cosh t \leq \cosh(\frac{2c}{\pi} x + d)$. Now, the lemma implies that if this inequality holds for all t then $a = \frac{2c}{\pi} \leq 1 \leq u = \frac{2c}{\pi}$. Thus $\frac{2c}{\pi} = 1$, but since $c = \frac{1}{2}(\lambda_2 - \lambda_1)$ we must have $\lambda_2 = \lambda_1 + \pi$. However, this is impossible since the corresponding LT $L_1(\lambda) = (\cos \lambda)^{-1}$ is not defined outside of the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$. To end up this discussion, no mixing can give increasing hazard rate for the NEF generated by the density $s_1(x) = \frac{1}{2\cosh \frac{\pi x}{2}}$.

The hyperbolic case $\alpha = 2$: As we are going to see here, the situation is more favorable when dealing with the following direct consequence of (14):

$$s_2(x) = \frac{x}{2 \sinh \frac{\pi x}{2}}.$$

Proposition 4 has shown that $b_2 = -\log s_2$ is a convex function, so we are in position to use Proposition 1. We have the explicit calculation

$$b_2''(x) = \frac{1}{x^2} - \left(\frac{\pi}{2}\right)^2 \frac{1}{\sinh^2 \frac{\pi x}{2}}.$$

Since $|t| \leq |\sinh t|$, clearly $b_2''(x) > 0$ and we have a direct proof of the log concavity of s_2. Thus we have to study the set of (c,d)'s such that the inequality

$$cT(x) = c\frac{x \sinh \frac{\pi x}{2}}{\sqrt{\sinh^2 \frac{\pi x}{2} - (\frac{\pi x}{2})^2}} \leq \cosh(cx+d) \tag{19}$$

holds for all real x. For this we use the following lemma.

Lemma 6 *For all real t we have*

$$\frac{t \sinh t}{\sqrt{\sinh^2 t - t^2}} \leq \sqrt{3+t^2},$$

where an equality occurs when $t = 0$.

Proof. The proof follows from the inequality $\sinh^2 t - t^2 - \frac{t^4}{3} \geq 0$, which is deduced from the expansion of

$$\sinh^2 t - t^2 - \frac{t^4}{3} = \frac{1}{2}\cosh 2t - \frac{1}{2} - t^2 - \frac{t^4}{3} = \sum_{n=3}^{\infty} \frac{2^{2n} t^{2n}}{(2n)!} \geq 0.$$

■

In terms of the function T, the lemma is equivalent to stating that for all x we have

$$T(x) \leq \frac{2}{\pi}\sqrt{3 + (\frac{\pi x}{2})^2}$$

Since it is difficult to find all (c,d) such that (9) holds we shall content to study the set of (c,d) such that $c\frac{2}{\pi}\sqrt{3+(\frac{\pi x}{2})^2} \leq \cosh(cx+d)$ holds for all x, or equivalently, by introducing $k = 2c/\pi$ and $u = k\frac{\pi x}{2} + d$, to study the set of (k,d) such that

$$\sqrt{3k^2 + (u-d)^2} \leq \cosh u$$

holds for all u.

Lemma 7 *For $k > 0$ and d real $\sqrt{3k^2+(u-d)^2} \leq \cosh u$ holds for all u if and only if $|d| \leq d_0 = \sqrt{2} - \log(1+\sqrt{2}) = 0.532...$ and*

$$3k^2 \leq (2 - \cosh^2 u_d)\cosh^2 u_d,$$

where u_d is the solution of the equation $\sinh 2u = 2(u-d)$. In particular, the inequality $\sqrt{3k^2+u^2} \leq \cosh u$ holds for all u if and only if $k \leq \sqrt{\frac{2}{3}}$.

Proof. The inequality $\sqrt{3k^2+(u-d)^2} \leq \cosh u$ implies $|u-d| \leq \cosh u$ for all u. Now the minimum d_0 of $\cosh u - u$ is attained at $\log(1+\sqrt{2})$ which is the solution of the equation $\sinh u - 1 = 0$, and thus, $d_0 = \sqrt{2} - \log(1+\sqrt{2})$.

Similarly, the minimum of $\cosh u + u$ is attained at $-\log(1+\sqrt{2})$ and is d_0. Since, $-\cosh u - u \leq -d_0 \leq d_0 \leq \cosh u - u$, we get that $|u-d| \leq \cosh u$ for all u if and only if $|d| \leq d_0$. Now, fix $d \in [-d_0, d_0]$, then for finding all k such that $3k^2 + (u-d)^2 \leq \cosh^2 u$, we look for the (positive) minimum of $u \mapsto \cosh^2 u - (u-d)^2$, which is attained at the point u_d. Letting $d=0$ gives $u_d = 0$ and also entails the final result. ■

Practical conclusion for $\alpha = 2$. The Laplace transform of s_2 is $1/\cos^2\lambda$ for $\lambda \in (-\frac{\pi}{2}, \frac{\pi}{2})$. According to Lemma 7, we fix any d such that $|d| \leq 0.532..$ and $c > 0$ such that $c \leq \frac{\pi}{2}\sqrt{\frac{2}{3}} = 1.282..$ We now choose an arbitrary number λ such that

$$-\frac{\pi}{2} < \lambda_1 = \lambda - c < \lambda + c = \lambda_2 < \frac{\pi}{2}$$

and we use formula (11) for defining the mixing coefficient p depending on d, λ_1 and λ_2. With this choice, the density

$$\frac{x}{2\sinh\frac{\pi x}{2}}[pe^{-\lambda_1 x}\cos^2\lambda_1 + (1-p)e^{-\lambda_2 x}\cos^2\lambda_2]$$

has an increasing hasard rate.

5 The Ressel NEF

Consider the density s_α on the positive real line defined by

$$s_\alpha(x) = e^{-b_\alpha(x)} = \frac{\alpha\, x^{x+\alpha-1}e^{-x}}{\Gamma(x+\alpha+1)},\ x > 0,\ \alpha > 0, \tag{20}$$

then s_α is called the Ressel or the Kendall-Ressel density with parameter α. It is infinitely divisible (in other terms the Jorgensen set $J(s_1)$ is $[0,\infty)$) and

$$s_\alpha * s_{\alpha'} = s_{\alpha+\alpha'}.$$

This density appears in various areas. For an M/G/1 queueing system with arrival rate λ, it is the limiting distribution, as $\lambda \to \infty$, of the length of the busy period $T(\alpha) - \alpha$ initiated by the virtual time quantity $\alpha > 0$ (c.f. Prabhu,1965, pages 73 and 237). In their characterization of the regression of the sample variance on the sample mean, Fosam and Shanbhag (1997) showed that such a regression is cubic on the sample mean for only for six distributions, of which one is the Kendall-Ressel distribution. Kokonendji (2001) also revealed this distribution on his investigation of first passage times on 0 and 1 of some Lévy processes for NEF's. Additional references can be provided here regarding the Kendall-Ressel distribution like Pakes (1995) formula (4.1), but one of the most detailed reference is Letac and Mora (1990) who characterized all NEF's having cubic VF's, of which, of course, the Ressel NEF is one of them.

The Ressel NEF generated by (20) has a VF $(V,\Omega) = \left(\frac{\mu^2}{\alpha}(1+\frac{\mu}{\alpha}), (0,\infty)\right)$. We are interested in the values of α such that $b''_\alpha(x) \geq 0$ for all $x > 0$. One

can consult Proposition 5.5 of Letac and Mora (1990) for checking the puzzling formula $\int_0^\infty s_\alpha(x)dx = 1$ and page 36 of this reference for learning why this density can also be called the Kendall-Ressel density.

Proposition 8 *Let*

$$g(x) = \frac{\alpha-1}{x^2} + \frac{(2-\alpha)x - \alpha^2 + \alpha}{x(x+\alpha)} \text{ and } h(x) = \frac{(\alpha^2-1) + (\alpha-\alpha^2)x + (2-\alpha)x^2}{x^2(x+\alpha+1)}, \tag{21}$$

then for all $x > 0$ we have $h(x) \leq b''_\alpha(x) \leq g(x)$. Furthermore there exists a number $a \in (1.77, 1.91)$ such that $b''_\alpha(x) \geq 0$ for all $x > 0$ if and only if $\alpha \in [1, a]$. In other terms the Glaser set $G(s_1)$ is $[1, a]$.

Proof. For $x > 0$, we use the digamma function $\psi = \Gamma'/\Gamma$ and formula (17). Using this notation we have

$$\begin{aligned} f(x) = b''_\alpha(x) &= -\frac{\alpha-1}{x} + \frac{\alpha-1}{x^2} + \psi'(x+\alpha+1) \\ &= -\frac{\alpha-1}{x} + \frac{\alpha-1}{x^2} + \sum_{n=2}^{\infty} \frac{1}{(n+x+\alpha-1)^2}. \end{aligned} \tag{22}$$

Now observe that

$$\frac{1}{1+x+\alpha} = \int_2^\infty \frac{dt}{(t+x+\alpha-1)^2} < \sum_{n=2}^{\infty} \frac{1}{(n+x+\alpha-1)^2} < \frac{1}{x+\alpha} = \int_1^\infty \frac{dt}{(t+x+\alpha-1)^2}.$$

This gives the desired inequalities $h \leq f \leq g$. Clearly the function f is positive if $\alpha = 1$. If $\alpha < 1$, the function f is equivalent in a neighborhood of $x = 0$ to $\frac{\alpha-1}{x^2}$ which tends to $-\infty$. It is obvious that for $\alpha \geq 2$, $g(x)$ becomes negative ultimately (if $\alpha = 2$ then $g(x) = (2-x)/x^2(2+x)$). Hence assume that $\alpha < 2$ and let us study the sign of g. Since $g(x) = 0$ if

$$(\alpha^2 - \alpha) + (-1 + 2\alpha - \alpha^2)x + (2-\alpha)x^2 = 0,$$

then this equation has at least one solution if

$$D(\alpha) = (-1 + 2\alpha + \alpha^2)^2 - 4(2-\alpha)(\alpha^2-\alpha) = 1 + 4\alpha - 6\alpha^2 + \alpha^4$$

is nonnegative, which is the case for $\alpha > \alpha^* = 1.90321$. One of the two possible solutions of $g(x) = 0$ is

$$x_1 = \frac{1 - 2\alpha + \alpha^2 + \sqrt{D(\alpha)}}{4-2\alpha} > \frac{(\alpha^*)^2 - 3}{4-2\alpha} > 0, \quad \alpha^* < \alpha < 2.$$

So $g(x_1) = 0$ and hence $g(x) < 0$ for some $x > 0$. On the other hand, we use inequality $h \leq f$ and study the sign of h. Let $1 < \alpha < 2$, then if $D(\alpha) = (\alpha-1)(-8 - 4\alpha + 3\alpha^2 + \alpha^3) < 0$, $h(x)$ has no zeros at all and hence $h(x) > 0$ for all $x > 0$. Now, $D(\alpha) = 0$ if $\alpha \in \{-3.48929, -1.28917, 1.0, 1.77846\}$ and

$D(\alpha) \to \infty$ as $\alpha \to \infty$, hence $D(\alpha) < 0$ (i.e., f is positive for all $x > 0$) if $1 < \alpha < 1.77846$. ■

Remarks. For studying the log concavity of the density s_α of the Ressel distribution one can be tempted to imitate Proposition 4 and to wonder if for $\alpha > 1$ the function

$$\frac{1}{s_\alpha(x)} = \frac{1}{\alpha}\frac{1}{x^{\alpha-1}}e^x\Gamma(x+\alpha+1) \times x^{-x}$$

defined on $(0, \infty)$ is the LT of a positive measure. Proposition 8 has shown that is is impossible if $\alpha > a$. This can be explained by the fact that the factor $x \mapsto x^{-x}$ is the only factor in $1/s_\alpha$ which is not a LT (this observation follows from the fact that $x \mapsto x^x$ is the LT of a stable law with parameter 1, and the reciprocal of the LT of a non Dirac measure cannot be a LT). In terms of the Glaser and Karlin sets, the density s_1 is quite interesting. Proposition 8 has shown that $G(s_1) = [1, a] \subset K(s_1)$ and a striking consequence is that while the density s_1 is log concave the density $s_2 = s_1 * s_1$ is not. This demonstrates the difference between the Glaser and the Karlin sets. The additive semigroup generated by $[1, a]$ is $[1, a] \cup [2, \infty) \subset K(s_1)$. One can reasonably conjecture that $K(s_1) = [1, \infty)$.

From Proposition 8 it follows that if $\alpha \in [1, a]$, we are led to consider $T(x) = 1/\sqrt{b''_\alpha(x)}$ and study the set of the pairs (c, d) such that $cT(x) \leq \cosh(cx + d)$. Since $h(x) \leq b''_\alpha(x) = -(s'_\alpha/s_\alpha)'(x)$, a sufficient condition for these (c, d) values is that $c/\sqrt{h(x)} \leq \cosh(cx + d)$. Even this simplified inequality is still too complicated and we shall content here to consider only the case $\alpha = 1$. For this case, we search for the set of (c, d)'s such that for all $x > 0$ we have $c\sqrt{x+2} \leq \cosh(cx + d)$. Therefore, the next proposition is devoted to the case of the NEF generated by the probability density

$$s_1(x) = \frac{x^x e^{-x}}{\Gamma(x+2)}\mathbf{1}_{(0,\infty)}(x)dx. \tag{23}$$

Proposition 9 *For $c > 0$ and d real consider the function*

$$\varphi(x) = \frac{1}{c^2}\cosh^2(cx + d) - x - 2$$

and define $x_0 = \frac{1}{c}\left(\frac{1}{2}\log(c + \sqrt{c^2+1}) - d\right)$. Then $\varphi(x) > 0$ for all $x > 0$ if and only if

1. either $x_0 \leq 0$ and $\cosh d \geq \sqrt{2}c$;

2. or $x_0 \geq 0$, $c \leq \sqrt{\frac{8}{7}}$ and $x_0 \leq \frac{1}{2c^2}(1 + \sqrt{c^2+1}) - \frac{7}{4}$.

Proof. Since $\varphi'(x) = \frac{1}{c}\sinh 2(cx+d) - 1$, $\varphi'(x_0) \leq 0$ for $x \leq x_0$ and $\varphi'(x_0) \geq 0$ for $x \geq x_0$, it follows that for $x_0 \leq 0$, $\varphi(x) \geq 0$ for $x \geq 0$ if and only if $\varphi(0) \geq 0$ or if and only if $\cosh d \geq \sqrt{2}c$ (recall that $x_0 \leq 0$ implies $d > 0$). Similarly

for $x_0 \geq 0$, we have $\varphi(x) \geq 0$ for $x \geq 0$ if and only if $\varphi(x_0) \geq 0$. Since $\frac{1}{c^2}\cosh^2(cx_0+d) = \frac{1}{2c^2}(1+\sqrt{c^2+1})$, the inequality $\varphi(x_0) \geq 0$ is equivalent to $x_0 \leq \frac{1}{2c^2}(1+\sqrt{c^2+1}) - \frac{7}{4}$, which can be realized only if the right hand side is nonnegative, that is if $c \leq \sqrt{\frac{8}{7}}$. ■

Now, having pairs (c,d) to our disposal, we have to compute p as given by the formula (11) and we need for this the values of the Laplace transform at points $\lambda_1 = \lambda - c$ and $\lambda_2 = \lambda + c$. For the Ressel distribution s_1, its LT

$$L_1(\lambda) = \int_0^\infty e^{-\lambda x} s_1(x)dx \tag{24}$$

cannot be expressed explicitly. However, a numerical or graphical calculation of $L_1(\lambda)$ for a given positive value of λ is easily done by means of the following proposition. Its statement is equivalent to formula (11) in Fosam and Shanbhag (1997) which relies on Prabhu (1965, p. 73 and p. 237). We give an independent proof here for sake of completeness.

Proposition 10 *Consider the bijection f from $[1,\infty)$ to itself defined by $f(x) = x - \log x$. Then for a given $\lambda > 0$, the number $1/L_1(\lambda)$ defined by (24) satisfies $f(1/L_1(\lambda)) = 1 + \lambda$.*

Proof. Let $(Y(t))_{t\geq 0}$ be the Lévy process such that for $t, \lambda \geq 0$ we have $\mathbb{E}(e^{-\lambda Y(t)}) = (1+\lambda)^{-t}$ (such a process is usually called the gamma process). Define $T = \min\{t : t - Y(t) = 1\}$. The random variable $T-1$ has the Ressel distribution (23) (see also Letac and Mora, 1990, p. 27 for this observation). Furthermore, Theorem 5.3 in Letac and Mora (1990) states that if $\mathbb{E}(e^{\theta T}) = e^{-q(\theta)}$ and $\mathbb{E}(e^{\theta(1-Y(1))}) = e^{-r(\theta)}$ then $r(q(\theta)) = \theta$. Thus, since $L_1(\lambda) = \mathbb{E}(e^{-\lambda(T-1)}) = e^{\lambda - q(-\lambda)}$ and since $e^{-r(\theta)} = \frac{e^\theta}{1+\theta}$, we can write $r(q(\theta)) = \log(1+q(\theta)) - q(\theta) = \theta$. This leads to

$$\log(1+q(-\lambda)) = q(-\lambda) - \lambda = -\log L_1(\lambda)$$

The elimination of $q(-\lambda)$ between these two equalities gives $\frac{1}{L_1(\lambda)} + \log L_1(\lambda) = 1 + \lambda$. Since, for $\lambda \geq 0$ we have $L_1(\lambda) \leq 1$, we get $f(1/L_1(\lambda)) = 1 + \lambda$. ■

Practical conclusion for $\alpha = 1$. We fix a pair (c,d) such that either condition 1 or condition 2 of Proposition 9 holds. We choose a number λ such that $0 < \lambda_1 = \lambda - c < \lambda + c = \lambda_2$. We compute numerically $L_1(\lambda_1)$ and $L_1(\lambda_2)$ with Proposition 10. The mixing coefficient p is therefore determined by (11). The density on $(0,\infty)$

$$\frac{x^x e^{-x}}{\Gamma(x+2)}\left[p\frac{e^{-\lambda_1 x}}{L_1(\lambda_1)} + (1-p)\frac{e^{-\lambda_2 x}}{L_1(\lambda_2)}\right]$$

has an increasing hazard rate.

6 The Kummer type 2 NEF

Let $a, \lambda > 0$ and b real and consider the number

$$C(a,b,\lambda) = \int_0^\infty \frac{x^{a-1}}{(1+x)^{a+b}} e^{-\lambda x} dx.$$

As a function of λ, C is proportional to what is sometimes called in the literature the confluent hypergeometric function of the second kind or a Whittaker function. If

$$s(x) = \frac{1}{C(a,b,\lambda)} \frac{x^{a-1}}{(1+x)^{a+b}} e^{-\lambda x} \mathbf{1}_{(0,\infty)}(x), \tag{25}$$

then the probability $s(x)dx = K^{(2)}(a,b,\lambda)(dx)$ is called the Kummer distribution of type 2 with parameters (a,b,λ). Needless to say, that if (a,b) are fixed the model $\{s(x)dx,\ \lambda > 0\}$ is an NEF. If $b > 0$, this model is generated by the beta distribution of type 2, i.e., by

$$\beta^{(2)}(a,b)(dx) = \frac{1}{B(a,b)} \frac{x^{a-1}}{(1+x)^{a+b}} \mathbf{1}_{(0,\infty)}(x) dx.$$

Kummer distributions have been studied by Ng and Kotz (1995). Statistical aspects of Kummer distributions for waiting times and exceedance statistics have been considered by Fitzgerald (2002). The Kummer distributions of type one belong to NEF's generated by the ordinary beta distributions. Since they are concentrated on the bounded set $(0,1)$, they are not relevant for our study. Accordingly, we study the NEF generated the Kummer distribution of type 2. Its VF cannot be expressed explicitly. However, the important fact about such an NEF is the formula (27) below which gives the LT of $\beta^{(2)}(a,b)$ in terms of the confluent hypergeometric function defined for real a and b such that b is not in the set $-\mathbb{N}$. This LT is then given in terms of the entire function

$$_1F_1(a\,;b\,;\lambda) = \sum_{n=0}^{\infty} \frac{(a)_n \lambda^n}{n!(b)_n}. \tag{26}$$

Here, $(a)_0 = 1$ and $(a)_{n+1} = (a+n)(a)_n$. This formula states that if $a > 0$ and b is not in the set $\mathbb{Z}$ of relative integers, we have

$$C(a,b,\lambda) = \frac{\Gamma(b)\Gamma(a)}{\Gamma(a+b)}\ {}_1F_1(a\,;1-b\,;\lambda) + \Gamma(-b)\lambda^b\ {}_1F_1(a+b\,;1+b\,;\lambda). \tag{27}$$

In (27), the mapping $z \mapsto 1/\Gamma(z)$ is an entire function which coincides with the ordinary $1/\Gamma(z), z > 0$, and (27) can be extended to the case where $b \in \mathbb{Z}$ by a limiting process. The identity (27) is by no means elementary and its proof by the Barnes formula can be found for instance in Slater (1960), formula 3.1.19. A probabilistic proof would be desirable.

To exemplify the use of (27), observe that if $a, b, \lambda > 0$ and if $X \sim \gamma(b,\lambda)$, $Y \sim K^{(2)}(a,b,\lambda)$ and $Z \sim \gamma(a+b,\lambda)$ are independent, then $X+Y$ and $Z/(1+Y)$

have the same distribution $K^{(2)}(a+b,-b,\lambda)$. In order to prove this, for suitable $t's$ just consider the LT $\mathbb{E}(e^{-t(X+Y)})$ and the Mellin transform $\mathbb{E}(Z^t/(1+Y)^t)$, and then use (27).

Proposition 11 *If s is defined by (25) then $-(s'/s)' > 0$ for all $x > 0$ if and only if $1 \leq a$ and $b \leq -1$ with $a-b-2 \neq 0$.*

Proof. If $A = a-1$ and $B = -b-1$, we get

$$-(\frac{s'}{s})'(x) = \frac{A + 2Ax + Bx^2}{x^2(1+x)^2}. \tag{28}$$

Trivially, $-(s'/s)' > 0$ if $A \geq 0$ and $B \geq 0$ with $A+B \neq 0$. Conversely, if $A + 2Ax + Bx^2 > 0$ for all $x > 0$, then letting $x \to \infty$ shows that $B \geq 0$. Also, letting $x \to 0$ shows that $A \geq 0$, while $A = B = 0$ would imply that $-(s'/s)' = 0$.Therefore if $A = a-1 \geq 0$ and $B = -b-1$ with $AB \neq 0$, we are

allowed, by using (28), to consider for $x > 0$,

$$T(x) = 1/\sqrt{-(\frac{s'}{s})'(x)} = \frac{x + x^2}{\sqrt{A + 2Ax + Bx^2}}.$$

We then have to investigate which numbers $c > 0$ and d real are such that for all $x > 0$, we have

$$c\frac{x + x^2}{\sqrt{A + 2Ax + Bx^2}} \leq \cosh(cx + d)$$

For simplicity, we are going to treat only the particular case $A = 0$, and therefore to study the NEF

$$\frac{1}{C(1, -B-1, \lambda)}(1+x)^B e^{-\lambda x},$$

where B is a fixed positive constant. For this particular case we are looking for the values of (c, d) with $c > 0$ such that for all $x > 0$ we have $\frac{\sqrt{B}}{c}\cosh(cx+d) - x - 1 \geq 0$. ∎

Proposition 12 *For $B, c > 0$ and d real, consider the function defined on $\mathbb{R}$ by*

$$\varphi(x) = \frac{\sqrt{B}}{c}\cosh(cx + d) - x - 1$$

and define $x_0 = \frac{1}{c}\left(\log\frac{1+\sqrt{B+1}}{\sqrt{B}} - d\right)$. Then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only

1. either $x_0 \leq 0$, $\sqrt{B} \leq c$ and $\cosh d \geq \frac{c}{\sqrt{B}}$;

2. or $x_0 \geq 0$, $c \leq \sqrt{B+1}$ and $x_0 \leq 1 - \frac{\sqrt{B+1}}{c}$.

Proof. We study the function φ in an elementary way: We get that $\varphi'(x) = \sqrt{B}\sinh(cx+d) - 1$ satisfies $\varphi'(x) \leq 0$ for $x \leq x_0$ and $\varphi'(x) \geq 0$ for $x \geq x_0$. If $x_0 \leq 0$ then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only $\varphi(0) = \frac{\sqrt{B}}{c}\cosh d - 1 \geq 0$ and this proves part 1. If $x_0 \geq 0$, then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only

$$\varphi(x_0) = \frac{\sqrt{B+1}}{c} - 1 - x_0 \geq 0$$

which proves part 2. ■

Here, again, in order to apply the results of this section to formula (11) we have to compute the values of the LT $C(1, -B-1, \lambda)$, which can also be seen as a truncated gamma function. If B is an integer, C is easily computed by the binomial formula

$$C(1, -B-1, \lambda) = \int_0^\infty (1+x)^B e^{-\lambda x} dx = \frac{B!}{\lambda^{B+1}} \sum_{n=0}^{B} \frac{\lambda^n}{n!}.$$

If $B > 0$ is not an integer, (27) gives

$$\begin{aligned} C(1, -B-1, \lambda) &= e^\lambda \int_1^\infty x^B e^{-\lambda x} dx = \frac{\Gamma(B+1)}{\lambda^{B+1}} e^\lambda - \frac{1}{B+1} \; {}_1F_1(1; 2+B; \lambda) \\ &= \frac{\Gamma(B+1)}{\lambda^{B+1}} \left[e^\lambda - \sum_{n=0}^\infty \frac{\lambda^{B+n+1}}{\Gamma(B+n+2)} \right], \end{aligned}$$

but then we have to rely on numerical analysis for computing the corresponding values of the confluent hypergeometric function

$${}_1F_1(1; B+2; \lambda) = 1 + \sum_{n=1}^\infty \frac{\lambda^n}{(B+2)\dots(B+n+1)}$$

and use (27). A good reference for such numerical consideration aspects can be found in Abad and Sesma (1995).

Practical conclusion for $a = 1, b = -1 - B$. We fix a pair (c, d) such that either condition 1 or condition 2 of Proposition 12 holds. We choose a number λ such that $0 < \lambda_1 = \lambda - c < \lambda + c = \lambda_2$. We compute numerically $C(1; -1-B, \lambda_1)$ and $C(1; -1-B, \lambda_2)$. The mixing coefficient p is therefore determined by (11). The density on $(0, \infty)$

$$(1+x)^B \left[p \frac{e^{-\lambda_1 x}}{C(1; -1-B, \lambda_1)} + (1-p) \frac{e^{-\lambda_2 x}}{C(1; -1-B, \lambda_2)} \right]$$

has an increasing hazard rate.

References

[1] Abad, J. and Sesma, J. (1995). Computation of the Regular Confluent Hypergeometric Function. *The Mathematica J.*, **5**, 74-76.

[2] Barlow, R.E. and Proschan, F. (1965). *Mathematical Theory of Reliability*, Wiley, New York.

[3] Ben Salah, N. and Masmoudi, A. (2010). The real powers of the convolution of a gamma distribution and a Bernoulli distribution *J. Theor. Probab.*, **23**, to appear.

[4] Block, H.W., Li, Y. and Savits, T.H. (2005). Initial and final behavior of failure rate functions for mixtures and systems, *J. Appl. Probab.*, **40**, 721-740.

[5] Block, H.W., Li, Y. and Savits, T.H. (2005). Mixtures of normal distributions: Modality and failure rate, *Statist. Probab. Lett.*, **74**, 253-264.

[6] Fitzgerald, D.L. (2002). Tricomi and Kummer functions in occurrence, waiting times and exceedance statistics, *Stoch. Environ. Res. Risk Assess.,* **16**, 207-214.

[7] Fosam, E.B. and Shanbhag, D.N. (1997). An extended Laha-Lukacs characterization result based on a regression property, *J. Statist. Plann. Inference,* **63**, 173-186.

[8] Glaser, R.E. (1980). Bathtub and related failure rate characterization, *J. Amer. Statist. Assoc.*, **75**, 667-672.

[9] Gradshteyn, I.S and Ryzhik, I.M (1980). *Table of integrals, series and products.* Academic Press, New York.

[10] Karlin, S. (1968). *Total Positivity*, volume 1, Stanford University Press, Stanford.

[11] Karlin, S. and Proshan, F. (1968). Polya Types Distributions of Convolutions, *Ann. Math. Statist.,* **31** 721-736.

[12] Kokonendji, C.C. (2001). First passage times on zero and one for natural exponential families, *Statist. Probab. Lett.*, **51**(3), 293-298.

[13] Letac, G. and Mora, M. (1990). Natural exponential families with cubic variance functions, *Ann. Statist.*, **18**, 1-37.

[14] Morris, C.N. (1982). Natural exponential families with quadratic variance functions, *Ann. Statist.*, **10**, 65-80.

[15] Ng, K.W. and Kotz, S. (1995). Kummer-Gamma and Kummer-Beta univariate and multivariate distributions, *Research report*, Department of Statistics, The University of Hong Kong, Hong Kong.

[16] Pakes, A.G. (1996). A hitting time for Lévy processes with applications to dams and branching processes, *Ann. Fac. Sc. Toulouse Math. (6)*, **5** (3) 521-544.

[17] Prabhu, N.U. (1965). *Queues and Inventories: A Study of Basic Stochastic Processes*, Wiley, New York.

[18] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders* (2nd Edition). Springer, New York.

[19] Sibuya, M. (2006). Applications of Hyperbolic Secant Distributions, *Jpn. J. Appl. Statist.*, **35** (1), 17-47.

[20] Shanbhag, D.N. (1979). Diagonality of the Bhattacharyya matrix as a characterization, *Theory Prob. Appl.*, **23**, 430-433.

[21] Slater, L. J. (1960). *Confluent Hypergeometric Functions.* Cambridge, England: Cambridge University Press.

[22] Smyth, G.K. (1994). A note on modelling cross correlations: Hyperbolic secant regression, *Biometrika*, **81**, 396-402.

[23] Zolotarev, V.M. (1967). On the M-divisibility of stable laws, *Theory Prob. Appl.*, **12**, 509-508.